

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 14, 2003, 00:49:36 (Search time 166.888 Seconds  
(Without alignments)  
1569.471 Million cell updates/sec)

Title: US-09-698-781-17  
Sequence: 1 TLPVLLFL 9

Scoring table: BIOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2054640 segs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+2.p2n.model -DEV-xml  
-O/cgn2\_1/USPTO.spool/US09698781/runat\_07032003\_083459\_5322/app.query.fasta\_1.654  
-DB-GenEmbl -OFMT-fastap -SUFFIX-p2n.rge -MINMATCH-0.1 -LOOPLC-0 -LOOPEXT-0  
-UNITS-bits -START-1 -END-1 -MATRIX-Biosum62 -TRANS-human40.cdi -LIST-45  
-DOCALLIGN-200 -THR.SCORE-pct -THR.MAX-100 -THR.MIN-0 -ALIGN-15 -MODE-LOCAL  
-OUTFMT-pio -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000  
-USER-US09698781 @CGN\_1\_1\_2566 @runat\_07032003\_083459\_5322 -NCPU-6 -ICPU-3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NES\_SCORES-0 -WAIT -LONGLOG -DEV.TIMEOUT-120  
-WARN\_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6 -FGAPEXT-7  
-YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : GenEmbl:\*  
1: gb\_da:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_ov:\*  
22: em\_or:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pla:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	2128	6 AX335634	AX335634 Sequence
2	44	100.0	2128	9 HSCRISP3G	X95240 H.sapiens m
3	44	100.0	2138	9 HSPG28	X94323 H.sapiens m
4	44	100.0	2144	6 AX127587	AX127587 Sequence
5	44	100.0	90901	9 HSA17120	AL121974 Human DNA
6	44	100.0	107625	2 AC103066	AC103066 Rattus no
7	44	100.0	116742	9 AC069197	AC069197 Homo sapi
8	44	100.0	136129	2 AC011050	AC011050 Homo sapi
9	44	100.0	151752	2 AC010779	AC010779 Homo sapi
10	44	100.0	152908	9 AC025842	AC025842 Homo sapi
11	44	100.0	153003	2 AC094798	AC094798 Homo sapi
12	44	100.0	161582	2 AC020628	AC020628 Homo sapi
13	44	100.0	177387	2 AC090046	AC090046 Mus muscu
14	44	100.0	229084	2 AC121977	AC121977 Mus muscu
15	44	97.7	58558	2 AC024305	AC024305 Homo sapi
16	43	97.7	165158	2 AC103184	AC103184 Rattus no
17	43	97.7	183386	2 AC098004	AC098004 Rattus no
18	43	97.7	185560	2 AC110982	AC110982 Rattus no
19	42	95.5	111701	2 AC130906	AC130906 Rattus no
20	42	95.5	123947	9 AC005045	AC005045 Homo sapi
21	42	95.5	175021	2 AC098138	AC098138 Rattus no
22	42	95.5	175744	2 AC109189	AC109189 Mus muscu
23	42	95.5	237727	2 AC098291	AC098291 Rattus no
24	42	95.5	337133	2 AC110384	AC110384 Rattus no
25	41	93.2	763	6 AX053775	AX053775 Sequence
26	41	93.2	1108	6 AX053654	AX053654 Sequence
27	41	93.2	1757	8 AF003981	AF003981 Stellaria
28	41	93.2	2378	9 HUMTB	L04282 Human CACCC
29	41	93.2	2602	9 AK096857	AK096857 Homo sapi
30	41	93.2	2723	9 HSM804929	AL833616 Homo sapi
31	41	93.2	2858	9 AF039019	AF039019 Homo sapi
32	41	93.2	2952	9 AF432210	AF432210 Homo sapi
33	41	93.2	2997	10 BC026819	BC026819 Mus muscu
34	41	93.2	3032	9 HSA236885	AJ236885 Homo sapi
35	41	93.2	3323	10 AB025250	AB025250 Mus muscu
36	41	93.2	3741	10 AB025258	AB025258 Mus muscu
37	41	93.2	36736	8 SPCC645	AL049498 S.pombe c
38	41	93.2	86156	9 HUAC002550	AC002550 Human Chr
39	41	93.2	94737	9 AL596327	AL596327 Human DNA
40	41	93.2	96067	2 AP000953	AP000953 Homo sapi
41	41	93.2	98031	2 AC105808	AC105808 Rattus no
42	41	93.2	104595	2 AC131227	AC131227 Rattus no
43	41	93.2	114396	9 HS1098F8	AL003457 Homo sapi
44	41	93.2	135837	9 AC004837	AC004837 Homo sapi
45	41	93.2	137247	2 AC015865	AC015865 Homo sapi

RESULT 1

## ALIGNMENTS

AX35634  
LOCUS AX35634 2128 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 6143 from Patent WO0194629.  
ACCESSION AX35634  
VERSION AX35634.1 GI:18126353  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,  
Hortigan, S., Soppet, D. R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 6143 13-DEC-2001;  
FEATURES  
source location/Qualifiers  
1. .2128  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 734 a 397 c 380 g 617 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 30.2 Length: 2128  
Score: 44.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-698-781-17 (1-9) x AX35634 (1-2128)  
Qy 1 ThirleupheprovalleuLeuPhelu 9  
Db 19 ACATTATCCAGTCGTGTTCTCTG 45  
RESULT 2  
LOCUS HSCRISP3G 2128 bp mRNA linear PRI 12-APR-1996  
DEFINITION H.sapiens mRNA for cysteine-rich secretory protein-3.  
ACCESSION X95240  
VERSION X95240.1 GI:1262818  
KEYWORDS CRISP-3 gene; cysteine-rich secretory protein-3.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Kratschmar, J., Haendler, B., Eberspaecher, U., Roosterman, D.,  
Donner, P. and Schleuning, W. D.  
TITLE The human cysteine-rich secretory protein (CRISP) family. Primary  
structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3  
JOURNAL Eur. J. Biochem. 236 (3), 827-836 (1996)  
MEDLINE 8665901  
PUBMED 2 (bases 1 to 2128)  
REFERENCE 2  
AUTHORS Haendler, B.  
TITLE Direct Submission  
JOURNAL Submitted (18-JAN-1996) B. Haendler, Schering AG, ICMB, S109/517,  
13342 Berlin, FRG  
FEATURES  
source location/Qualifiers  
1. .2128  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="testis cDNA library"  
16. .753  
/gene="CRISP-3"  
16. .753  
/gene="CRISP-3"  
/codon\_start=1  
/product="cysteine-rich secretory protein-3"

/protein\_id="CAA64527.1"  
/db\_xref="GI:1262819"  
/db\_xref="SWISS-PROT:P54108"  
/translation="MTLFPVLLFLVAGLIPSPFANEDKDPATFALLTQTOVOREIVN  
KHNELEKRAVSPPARNNLKMENKEAANAKOAKNOCHYRHSNPKDRTSLKCGENILM  
SSASSSMSQAIOQSWFDEYNDPVGVPKTPNAVGHGTQVWVSYLVGCGNAYCPNQ  
KVLKYYVCOYCPAGMANRLVYVEGAGACASCPCNDGDLCTNGCKYEDLYSNCKS  
LKLTLCCKHQLVDRSCASCNSIY"  
16. .75  
/gene="CRISP-3"  
polyA.signal 2084..2089  
BASE COUNT 734 a 397 c 380 g 617 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 30.2 Length: 2128  
Score: 44.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-698-781-17 (1-9) x HSCRISP3G (1-2128)  
Qy 1 ThirleupheprovalleuLeuPhelu 9  
Db 19 ACATTATCCAGTCGTGTTCTCTG 45  
RESULT 3  
LOCUS HSSPG28 2138 bp mRNA linear PRI 04-MAR-1996  
DEFINITION H.sapiens mRNA for SGP28 protein.  
ACCESSION X94323  
VERSION X94323.1 GI:1213612  
KEYWORDS glycoprotein; SGP28 protein.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Kjeldsen, L., Cowland, J. B., Johnsen, A. H. and Borregaard, N.  
TITLE SGP28, a novel matrix glycoprotein in specific granules of human  
neutrophils with similarity to a human testis-specific gene product  
and a rodent sperm-coating glycoprotein  
JOURNAL FEBS Lett. 380 (3), 246-250 (1996)  
MEDLINE 96186934  
PUBMED 2 (bases 1 to 2138)  
REFERENCE 2  
AUTHORS Cowland, J. B.  
TITLE Direct Submission  
JOURNAL Submitted (19-DEC-1995) J. B. Cowland, Granulocyte Research  
Laboratory, Dept of Hematology, National Univ. Hosp.,  
Rigshospitalet L-4041, 9 Blegdamsvej, 2100 Copenhagen, DENMARK  
FEATURES  
source location/Qualifiers  
1. .2138  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="bone marrow"  
41. .778  
/codon\_start=1  
/product="SGP28 protein"  
/protein\_id="CAA63984.1"  
/db\_xref="GI:1213613"  
/translation="MTLFPVLLFLVAGLIPSPFANEDKDPATFALLTQTOVOREIVN  
KHNELEKRAVSPPARNNLKMENKEAANAKOAKNOCHYRHSNPKDRTSLKCGENILM  
SSASSSMSQAIOQSWFDEYNDPVGVPKTPNAVGHGTQVWVSYLVGCGNAYCPNQ  
KVLKYYVCOYCPAGMANRLVYVEGAGACASCPCNDGDLCTNGCKYEDLYSNCKS  
LKLTLCCKHQLVDRSCASCNSIY"  
BASE COUNT 728 a 404 c 383 g 623 t  
ORIGIN  
Alignment Scores:



repeat\_region /note="MER47 repeat: matches 2060. ,2246 of consensus"  
32920. ,33131  
/note="MER47 repeat: matches 1. ,218 of consensus"  
33336. ,33414  
repeat\_region /note="MER5B repeat: matches 68. ,155 of consensus"  
33550. ,33593  
repeat\_region /note="22 copies 2 mer ta 93% conserved"  
35966. ,36808  
repeat\_region /note="L1PA2 repeat: matches 5304. ,6146 of consensus"  
36808. ,36987  
repeat\_region /note="L1P repeat: matches 4908. ,5087 of consensus"  
37363. ,37490  
repeat\_region /note="L2 repeat: matches 2284. ,2414 of consensus"  
37542. ,37629  
misc\_feature /note="L2 repeat: matches 2421. ,2508 of consensus"  
complement(37721. ,38052)  
/note="match: GSS: Em:AD178503"  
38228. ,38316  
repeat\_region /note="MADE1 repeat: matches 1. ,80 of consensus"  
39513. ,39548  
repeat\_region /note="18 copies 2 mer ac 97% conserved"  
41609. ,41686  
repeat\_region /note="39 copies 2 mer tt 69% conserved"  
41687. ,42177  
repeat\_region /note="L1PA4 repeat: matches 5654. ,6144 of consensus"  
42678. ,43108  
repeat\_region /note="L1R19A repeat: matches 1. ,411 of consensus"  
43437. ,44028  
repeat\_region /note="Charlie2 repeat: matches 264. ,875 of consensus"  
45514. ,45903  
repeat\_region /note="MER93 repeat: matches 13. ,397 of consensus"  
46581. ,46898  
repeat\_region /note="L1R33 repeat: matches 4. ,356 of consensus"  
47008. ,47065

Alignment Scores:  
Pred. No.:  
Score: 1.28e+03 Length: 90901  
Percent Similarity: 44.00 Matches: 9  
Percent Local Similarity: 100.00% Conservative: 0  
Query Match: 100.00% Mismatches: 0  
DB: 9 Indels: 0  
Gaps: 0

US-09-698-781-17 (1-9) x HSJ417L20 (1-90901)  
QY 1 ThirleupheProvalleuThirleupheleu 9  
Db 7851 ACATTATTCCACGTCCTTCTTCCTG 7825

RESULT 6  
AC103066/c 107625 bp DNA linear HTG 13-JUL-2002  
LOCUS Rattus norvegicus clone CH230-170F5, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION \*\*\* 68 unordered pieces.  
AC103066 GI:21730986  
AC103066.3 GI:21730986  
VERSION HTG; HTGS\_PHASEL.  
KEYWORDS Norway rat.  
SOURCE Rattus norvegicus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
1 (bases 1 to 107625)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,  
Aisbrooks,S.L., Amaratunga,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonini,D.,  
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,K., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Devalla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dlinh,H.H.,

```

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferriguto,D., Flaag,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homsel,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsom,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
L.J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peterson,L., Pickens,R., Pitman,E., Pu,L.L., Quties,M., Ren,X.,
Rivers,M., Rojas,A., Rojibokan,I., Roite,M., Ruiz,S., Saverly,G.,
Scheer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usman,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
2 (bases 1 to 107625)
Worley,K.C.
Direct Submission
Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 107625)
Worley,K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:17974481.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center Project name: GUDG
Center Clone name: CH230-170P5
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 29809 bases at least Q40
Consensus quality: 34249 bases at least Q30
Consensus quality: 36749 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 68 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1141: contig of 1141 bp in length
1142 1241: gap of unknown length

```

```

1242 2424: contig of 1183 bp in length
2425 2524: gap of unknown length
2525 3588: contig of 1064 bp in length
3589 3688: gap of unknown length
3689 5276: contig of 1588 bp in length
5277 5376: gap of unknown length
5377 6534: contig of 1158 bp in length
6535 7795: gap of unknown length
7796 7959: contig of 1161 bp in length
7960 8986: gap of unknown length
8987 9086: contig of 1091 bp in length
9087 10789: gap of unknown length
10790 10889: contig of 1703 bp in length
10890 11971: gap of unknown length
11972 12071: contig of 1082 bp in length
12072 13077: gap of unknown length
13078 13177: contig of 1006 bp in length
13178 14680: gap of unknown length
14681 14780: contig of 1503 bp in length
14781 15868: gap of unknown length
15869 17569: contig of 1088 bp in length
17570 17669: gap of unknown length
17670 19184: contig of 1601 bp in length
19185 19284: gap of unknown length
19285 20795: contig of 1515 bp in length
20796 20895: gap of unknown length
20896 22456: contig of 1511 bp in length
22457 22556: gap of unknown length
22557 23587: contig of 1030 bp in length
23588 24797: gap of unknown length
24798 24897: contig of 1111 bp in length
24898 25936: gap of unknown length
25937 26037: contig of 1039 bp in length
26038 27468: gap of unknown length
27469 27568: contig of 1432 bp in length
27569 28991: gap of unknown length
28992 30264: contig of 1423 bp in length
30265 30364: gap of unknown length
30365 31609: contig of 1173 bp in length
31610 33499: gap of unknown length
33500 33599: contig of 1790 bp in length
33600 35051: gap of unknown length
35052 35151: contig of 1452 bp in length
35152 36896: gap of unknown length
36897 38705: contig of 1745 bp in length
38706 38805: gap of unknown length
38806 40482: contig of 1709 bp in length
40483 40582: gap of unknown length
40583 42236: contig of 1677 bp in length
42237 42336: gap of unknown length
42337 43988: contig of 1654 bp in length
43989 44088: gap of unknown length
44089 45246: contig of 1552 bp in length
45247 45347: gap of unknown length
45348 46935: contig of 1158 bp in length
46936 47035: gap of unknown length
47036 48169: contig of 1589 bp in length
48170 4837: gap of unknown length
4838 49437: contig of 1034 bp in length
49439 49537: gap of unknown length
49538 51697: contig of 1268 bp in length
51698 51797: gap of unknown length
51799 53050: contig of 2160 bp in length
53051 53150: gap of unknown length
53151 54318: contig of 1253 bp in length
54319 55908: gap of unknown length
55909 55908: contig of 1168 bp in length
55910 55908: gap of unknown length
55911 55908: contig of 1490 bp in length

```



```

repeat_region      2043. .2334      /rpt_family="AluSc"
repeat_region      complement(2335. .2622)
repeat_region      /rpt_family="L1MDa"
repeat_region      2673. .3064      /rpt_family="MLT1C"
repeat_region      complement(3070. .3928)
repeat_region      /rpt_family="L1MDa"
unsure
repeat_region      /note="<30 qual SINGL region"
repeat_region      4823. .4862      /rpt_family="AT_r1ch"
repeat_region      complement(4868. .4937)
repeat_region      /rpt_family="L2"
repeat_region      4971. .5169      /rpt_family="MIR"
repeat_region      5448. .5663      /rpt_family="LTR33"
repeat_region      5712. .5807      /rpt_family="LTR33"
repeat_region      6732. .6849      /rpt_family="LTR33"
repeat_region      /rpt_family="MER91C"
repeat_region      6880. .7126      /rpt_family="MIR"
repeat_region      complement(7348. .7438)
repeat_region      /rpt_family="L2"
repeat_region      complement(7711. .7958)
repeat_region      /rpt_family="L2"
repeat_region      complement(8333. .8626)
repeat_region      /rpt_family="AluSg"
repeat_region      complement(8680. .8993)
repeat_region      /rpt_family="AluY"
repeat_region      10527. .10615      /rpt_family="A-rich"
repeat_region      10743. .10826      /rpt_family="MER20"
repeat_region      10843. .11142      /rpt_family="AluSg"
repeat_region      11184. .11629      /rpt_family="MLT1C"
repeat_region      11630. .11731      /rpt_family="MER20"
repeat_region      11929. .11970      /rpt_family="L2"
repeat_region      12203. .12344      /rpt_family="MIR"
repeat_region      complement(12498. .12621)
repeat_region      /rpt_family="MIR"
repeat_region      complement(13416. .13769)
repeat_region      /rpt_family="MSTP"
repeat_region      complement(14815. .15129)
repeat_region      /rpt_family="AluSx"
repeat_region      17010. .17567      /rpt_family="MLT1E2"
repeat_region      18680. .18706      /rpt_family="TG1n"
repeat_region      complement(19182. .19248)
repeat_region      /rpt_family="L3"
repeat_region      19404. .19745      /rpt_family="L3"
repeat_region      complement(20225. .24814)
repeat_region      /rpt_family="L1MA9"
repeat_region      complement(24896. .25634)
repeat_region      /rpt_family="L1MA9"
repeat_region      25782. .26394      /rpt_family="L2"
repeat_region      complement(26601. .26690)
repeat_region      /rpt_family="MIR"
repeat_region      27295. .27425      /rpt_family="TA)n"
repeat_region      complement(28662. .28886)
repeat_region      /rpt_family="MIR"
repeat_region      29104. .29223

```

```

repeat_region      /rpt_family="MIR3"
repeat_region      complement(29424. .29992)
repeat_region      /rpt_family="MLT1K"
repeat_region      30757. .30777      /rpt_family="AT_r1ch"
repeat_region      30867. .31311      /rpt_family="L1MD"
repeat_region      32424. .32529      /rpt_family="L1MD"
repeat_region      complement(32636. .32928)
repeat_region      /rpt_family="AluSx"
repeat_region      complement(33257. .33695)
repeat_region      /rpt_family="L2"
repeat_region      33791. .33886      /rpt_family="(TA)n"
repeat_region      34742. .34888      /rpt_family="CT_r1ch"
repeat_region      36048. .36287      /rpt_family="MIR"
repeat_region      /rpt_family="MIR"
repeat_region      complement(36421. .36530)
repeat_region      /rpt_family="MIR"
repeat_region      37076. .37322      /rpt_family="MIR"
repeat_region      complement(37499. .37857)
repeat_region      /rpt_family="THE1C"
repeat_region      37927. .37992      /rpt_family="L2"
repeat_region      complement(38657. .39210)
repeat_region      /rpt_family="MER7"
repeat_region      39564. .39587

```

## Alignment Scores:

```

Pred. No.:      1.64e+03      Length:      116742
Score:      44.00      Matches:      9
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      9      Gaps:      0

```

US-09-698-781-17 (1-9) x AC069197 (1-116742)

QY 1 ThrleupheProValleupheleu 9

Db 19331 ACCTGTTCGCGTTTCTGTTTGTG 19357

## RESULT 8

AC011050/c

LOCUS AC011050 136129 bp DNA linear HTG 20-APR-2000

DEFINITION Homo sapiens clone Rpl1-9P22, WORKING DRAFT SEQUENCE, 14 unordered

pieces.

ACCESSION AC011050

VERSION AC011050.4

KEYWORDS GI:7596799

SOURCE HTG: HTGS\_PHASE1; HTGS\_DRAFT.

ORGANISM Homo sapiens.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 136129)  
 Homo sapiens, clone Rpl1-9P22  
 Unpublished  
 2 (bases 1 to 136129)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, D., Barna, N., Beckerly, R., Boguslavsky, L., Bouckgatter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hago, B., Heaford, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehocsky, J., Liu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Melidrim, J., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Peterson, K., Pollaro, V., Riley, R., Roy, A., Santos, R., Severy, P.,

TITLE  
JOURNAL

## COMMENT

Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Teste, S., Tlreil, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (30-SEP-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 19, 2000 this sequence version replaced gl:7107894.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: MIBR  
 Web site: http://www.seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L3027  
 Center clone name: 9\_P22

Summary Statistics  
 Sequencing vector: M13; M7815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 127628 bases at least Q40  
 Consensus quality: 131421 bases at least Q30  
 Consensus quality: 133129 bases at least Q20  
 Insert size: 150000; agarose-fp  
 Insert size: 134829; sum-of-coverage  
 Quality coverage: 3.9 in Q20 bases; agarose-fp  
 Quality coverage: 4.3 in Q20 bases; sum-of-coverage

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 14 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1581: contig of 1581 bp in length  
 \* 1582 1681: gap of 100 bp  
 \* 1682 3889: contig of 2208 bp in length  
 \* 3890 3989: gap of 100 bp  
 \* 3990 8412: contig of 4423 bp in length  
 \* 8413 8512: gap of 100 bp  
 \* 8513 12662: contig of 4150 bp in length  
 \* 12663 12762: gap of 100 bp  
 \* 12763 18164: contig of 5402 bp in length  
 \* 18165 18264: gap of 100 bp  
 \* 18265 22997: contig of 4733 bp in length  
 \* 22998 23087: gap of 100 bp  
 \* 23098 31525: contig of 8428 bp in length  
 \* 31526 31625: gap of 100 bp  
 \* 31626 41605: contig of 9980 bp in length  
 \* 41606 41705: gap of 100 bp  
 \* 41706 53440: contig of 11735 bp in length  
 \* 53441 53540: gap of 100 bp  
 \* 53541 63761: contig of 10221 bp in length  
 \* 63762 63861: gap of 100 bp  
 \* 63862 80106: contig of 16245 bp in length  
 \* 80107 80206: gap of 100 bp  
 \* 80207 95595: contig of 15389 bp in length  
 \* 95596 95695: gap of 100 bp  
 \* 95696 110749: contig of 15054 bp in length  
 \* 110750 110849: gap of 100 bp  
 \* 110850 136129: contig of 25280 bp in length.

FEATURES  
source

1. 136129  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="RP11-9P22"  
 /clone\_lib="RP11-11 Human Male BAC"  
 1. 1581  
 /note="assembly\_fragment"

misc\_feature 1682..3889  
 /note="assembly\_fragment"  
 misc\_feature 3990..8412  
 /note="assembly\_fragment"  
 misc\_feature 8513..12662  
 /note="assembly\_fragment"  
 misc\_feature 12763..18164  
 /note="assembly\_fragment"  
 misc\_feature 18265..22997  
 /note="assembly\_fragment"  
 misc\_feature 23098..31525  
 /note="assembly\_fragment"  
 misc\_feature 31626..41605  
 /note="assembly\_fragment"  
 misc\_feature 41706..53440  
 /note="assembly\_fragment"  
 misc\_feature 53541..63761  
 /note="assembly\_fragment"  
 misc\_feature 63862..80106  
 /note="assembly\_fragment"  
 misc\_feature 80207..95595  
 /note="assembly\_fragment"  
 misc\_feature 95696..110749  
 /note="assembly\_fragment"  
 misc\_feature 110850..136129  
 /note="assembly\_fragment"

BASE COUNT 43332 a 26887 c 26108 g 38701 t 1301 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 1.91e+03 length: 136129  
 Score: 44.00 Matches: 9  
 Percent Similarity: 100.008 Conservative: 0  
 Best Local Similarity: 100.008 Mismatches: 0  
 Query Match: 100.008 Indels: 0  
 DB: 2 Gaps: 0

US-09-698-781-17 (1-9) x AC011050 (1-136129)

OY 1 Thirleupheprovalleupheleu 9

Db 20656 ACCTGTTCGCGTTGTGTGTTTG 20630

## RESULT 9

AC010779/c AC010779 151752 bp DNA linear HTG 22-JUL-2000

LOCUS Homo sapiens clone RP11-3A4, WORKING DRAFT SEQUENCE, 11 unordered

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## TITLES

## JOURNAL

## AUTHORS

1 (bases 1 to 151752)  
 Homo sapiens  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
 Baldin, B., Barna, N., Becker, R., Boguslavsky, L., Boukhgalter, B.,  
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,  
 Cooke, P., Dearlano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,  
 Ferraira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,  
 Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
 Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
 Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,  
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,



TITLE	JOURNAL	COMMENT
Directed Submissions	Whitehead Institute/MIT Center for Genome Research, 350 Charles Street, Cambridge, MA 02141, USA	On Jul 22, 2000 this sequence version replaced g1:37521560.
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talmas, J., Testaye, S., Tirrell, A., Vassiliou, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.		

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.vi.mit.edu>  
 Contact: [sequence.submissions@genome.vi.mit.edu](mailto:sequence.submissions@genome.vi.mit.edu)  
 Project Information  
 Center Project name: I2699  
 Center Clone name: 3\_A4  
 Summary Statistics

NOTE: This is a working draft sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

*	1	1901:	contlg of	1901 bp	in	length
*	1902	2001:	gap of	100 bp	in	length
*	2002	4631:	contlg of	2630 bp	in	length
*	4632	4731:	gap of	100 bp	in	length
*	4732	9746:	contlg of	5015 bp	in	length
*	9747	9846:	gap of	100 bp	in	length
*	9847	17659:	contlg of	7823 bp	in	length
*	17670	17769:	gap of	100 bp	in	length
*	17770	29964:	contlg of	12195 bp	in	length
*	29965	30064:	gap of	100 bp	in	length
*	30065	43497:	contlg of	13433 bp	in	length
*	43498	43597:	gap of	100 bp	in	length
*	43598	60340:	contlg of	16743 bp	in	length
*	60341	60440:	gap of	100 bp	in	length
*	60441	75774:	contlg of	15534 bp	in	length
*	75775	75874:	gap of	100 bp	in	length
*	75875	93175:	contlg of	17301 bp	in	length
*	93176	93275:	gap of	100 bp	in	length
*	93276	119033:	contlg of	25764 bp	in	length
*	119040	119139:	gap of	100 bp	in	length
*	119140	151752:	contlg of	32613 bp	in	length

## FEATURES

1. 151/52

/organism="Homo sapiens"

```
/db_xref="taxon:9606"
```

```
clone="RP11-3A4"
```

/clone\_11b="RPC1-11 Human Male BAC

T06Y : . T

```

/note=assembly_fragment
2003  4631

```

2002. 14031  
TC04. 12007

71106e- assembly-fragment  
4732 97A6

```
4152: .3140
/notes="assembly" fragment"
```

/noce- uoocaiwty--ttaymenc

```
misc_feature      9847.  .17669  
/note="assembly-fragment"  
misc_feature      17770.  .29964  
/note="assembly-fragment"  
misc_feature      30065.  .43497  
/note="assembly-fragment"  
misc_feature      43598.  .60340  
/note="assembly-fragment  
clone_end:T7  
vector_side:right"  
60441.  .75774  
/note="assembly-fragment"  
misc_feature      75815.  .93175  
/note="assembly-fragment"  
misc_feature      93376.  .119039  
/note="assembly-fragment"  
misc_feature      119140.  .151752  
/note="assembly-fragment  
clone_end:SP6  
vector_side:right"
```

Alignment Scores:

Pred. No.:	2.13e+03	length:	151752
Score:	44.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-698-781-17 (1-9) x AC010779 (1-151752)

1 ThrLeupheprovalLeuleupheleu s

Db 88640 ACATTATTCCAGTGTCTGTTCTC 88614

RESULT	10			
AC025842/c				
LOCUS	AC025842	152908 bp	DNA	linear
DEFINITION	Homo sapiens chromosome 15, clone RP11-31E22, complete sequence.			
ACCESSION	AC025842			
VERSION	AC025842.11	GI:18482284		
KEYWORDS	HTG.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			

REFERENCE	1 (bases 1 to 152908)
AUTHORS	Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE	Homo sapiens chromosome 15, clone RP11-31E22
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 152908)
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

```

TITLE      Young,G., Zainoun,J., Zimmer,A. and Zody,M.
JOURNAL    Direct Submission
REFERENCE  Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome
AUTHORS    Research, 320 Charles Street, Cambridge, MA 02141, USA
            3 (bases 1 to 152908)

Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barta,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campilano,A., Chang,J., Chazaro,B.,
Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearliano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Menus,L., Mihova,T.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Roselli,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (15-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 3, 2002 this sequence version replaced g1:18450159.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center Project name: L6112
Center Clon name: 31_E_22
-----

FEATURES
source      1. .152908
             /location/Qualifiers
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="15"
             /map="15"
             /clone="RP11-31E22"
             /clone_lib="RPCI-11 Human Male BAC"
             /rpt_family="MLT1K"
             complement(409..528)
repeat_region
             /rpt_family="MIR3"
             746..970
             /rpt_family="MIR"
             2209..2263
             /rpt_family="(TA)n"
             2269..2359
             /rpt_family="LIMCa"
             2946..3035
             /rpt_family="MIR"
             complement(3242..3854)
repeat_region
             /rpt_family="L2"
             4002..4740
             /rpt_family="L1MA9"
             4822..5119
             /rpt_family="L1MA9"
             5103..9417
             /rpt_family="L1MA9"
             complement(9897..10238)
repeat_region
             /rpt_family="L3"

repeat_region 10394..10460
               /rpt_family="MIR"
repeat_region 10936..10962
               /rpt_family="(CA)n"
               complement(12075..12632)
repeat_region 14513..14827
               /rpt_family="MLT1E2"
repeat_region 15873..16226
               /rpt_family="Alusx"
               /rpt_family="MSRD"
               17021..17144
               /rpt_family="MIR"
               complement(17298..17439)
               /rpt_family="MIR"
               complement(17672..17713)
               /rpt_family="L2"
               complement(17911..18012)
               /rpt_family="MER20"
               complement(18013..18458)
               /rpt_family="MLT1C"
               complement(18500..18800)
               /rpt_family="Alusg"
               complement(18817..18900)
               /rpt_family="MER20"
               19028..19116
               /rpt_family="T-rich"
               20650..20963
               /rpt_family="Aluy"
               21017..21310
               /rpt_family="Alusg"
               21685..21932
               /rpt_family="L2"
               22205..22295
               /rpt_family="L2"
               complement(22517..22763)
               /rpt_family="MIR"
               22794..22911
               /rpt_family="MER91C"
               complement(23836..24208)
               /rpt_family="LTR33"
               complement(24487..24685)
               /rpt_family="MIR"
               24719..24788
               /rpt_family="L2"
               24794..24833
               /rpt_family="AT-rich"
               25728..26583
               /rpt_family="LIMDa"
               complement(26589..26980)
               /rpt_family="MLT1C"
               27032..27319
               /rpt_family="LIMDa"
               complement(27320..27611)
               /rpt_family="AlusC"
               27612..28205
               /rpt_family="LIMDa"
               28209..28511
               /rpt_family="LIM4"
               28589..29377
               /rpt_family="L2"
               complement(29431..29911)
               /rpt_family="L1PA10"
               29915..29963
               /rpt_family="CT-rich"
               complement(31398..31502)
               /rpt_family="MIR3"
               complement(31738..31860)
               /rpt_family="LIMEd"
               31861..32156
               /rpt_family="AlusC"
               complement(32157..32367)
               /rpt_family="LIMEd"
               complement(32443..34629)

```



```

* 37660 37759: gap of unknown length
* 37760 40181: contig of 2422 bp in length
* 40182 40281: gap of unknown length
* 40282 41668: contig of 1387 bp in length
* 41669 41768: gap of unknown length
* 41769 44738: contig of 2970 bp in length
* 44739 44838: gap of unknown length
* 44839 47294: contig of 2456 bp in length
* 47295 47394: gap of unknown length
* 47395 50279: contig of 2885 bp in length
* 50280 50379: gap of unknown length
* 50380 53411: contig of 3032 bp in length
* 53412 53511: gap of unknown length
* 53512 56852: contig of 3341 bp in length
* 56853 56952: gap of unknown length
* 56953 59456: contig of 2504 bp in length
* 59457 59556: gap of unknown length
* 59557 62291: contig of 2735 bp in length
* 62292 62391: gap of unknown length
* 62392 65871: contig of 3480 bp in length
* 65872 65971: gap of unknown length
* 65972 68691: contig of 2720 bp in length
* 68692 68791: gap of unknown length
* 68792 70548: contig of 1757 bp in length
* 70549 70648: gap of unknown length
* 70649 72765: contig of 2117 bp in length
* 72766 72865: gap of unknown length
* 72866 76266: contig of 3401 bp in length
* 76267 76366: gap of unknown length
* 76367 79074: contig of 2708 bp in length
* 79075 79174: gap of unknown length
* 79175 82058: contig of 2884 bp in length
* 82059 82158: gap of unknown length
* 82159 83557: contig of 1399 bp in length
* 83558 83657: gap of unknown length
* 83658 85936: contig of 2279 bp in length
* 85937 86036: gap of unknown length
* 86037 87516: contig of 1480 bp in length
* 87517 87616: gap of unknown length
* 87617 90413: contig of 2797 bp in length
* 90414 90513: gap of unknown length
* 90514 93572: contig of 3059 bp in length
* 93573 93672: gap of unknown length
* 93673 95740: contig of 2068 bp in length
* 95741 95840: gap of unknown length
* 95841 97292: contig of 1452 bp in length
* 97293 97392: gap of unknown length
* 97393 98500: contig of 1108 bp in length
* 98501 98600: gap of unknown length
* 98601 98994: contig of 1294 bp in length
* 98995 99994: gap of unknown length
* 99995 102248: contig of 2254 bp in length
* 102249 102348: gap of unknown length
* 102349 104341: contig of 2193 bp in length
* 104342 104641: gap of unknown length
* 104642 106217: contig of 1576 bp in length
* 106218 106317: gap of unknown length
* 106318 108044: contig of 1726 bp in length
* 108045 108143: gap of unknown length
* 108144 109955: contig of 1812 bp in length
* 109956 110055: gap of unknown length
* 110056 111732: contig of 1677 bp in length
* 111733 111832: gap of unknown length
* 111833 113204: contig of 1372 bp in length
* 113205 113304: gap of unknown length
* 113305 115196: contig of 1892 bp in length
* 115197 115296: gap of unknown length
* 115297 117372: contig of 2076 bp in length
* 117373 117472: gap of unknown length
* 117473 120048: contig of 2576 bp in length
* 120049 120148: gap of unknown length
* 120149 122391: contig of 2243 bp in length
* 122392 122491: gap of unknown length

```

```

* 122492 124665: contig of 2174 bp in length
* 124666 124765: gap of unknown length
* 124766 127338: contig of 2573 bp in length
* 127339 127438: gap of unknown length
* 127439 128567: contig of 1129 bp in length
* 128568 128667: gap of unknown length
* 128668 128922: contig of 1255 bp in length
* 128923 130022: gap of unknown length
* 130023 131375: contig of 1353 bp in length
* 131376 131475: gap of unknown length
* 131476 133004: contig of 1529 bp in length
* 133005 133104: gap of unknown length
* 133105 134605: contig of 1500 bp in length
* 134605 134704: gap of unknown length
* 134705 136391: contig of 1687 bp in length
* 136392 136491: gap of unknown length
* 136492 137494: contig of 1003 bp in length
* 137495 137594: gap of unknown length

Alignment Scores:
Pred. No.: 2.15e+03 Length: 153003
Score: 44.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-698-781-17 (1-9) x AC094798 (1-153003)
Qy 1 Thrieuphepovallenuleupheleu 9
Db 116823 ACCCTTTCCGTCCTCTTCTG 116849

RESULT 12
AC020628
LOCUS
DEFINITION
Homo sapiens chromosome 12 clone RP11-328H16, WORKING DRAFT
SEQUENCE, 7 unordered pieces.
AC020628
AC020628.16 GI:14647255
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
SOURCE
Homo sapiens.
ORGANISM
REFERENCE
AUTHORS
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-ouman,F.R., Allen,C.,
Alstrooms,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbarta,J.,
Benton,J., Bimaye,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gotrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,
Hollins,B., Homsi,F., Howard,S., Huber,J., Huliyil,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,D., Liu,M.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mathew,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabadi,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenko,S.,

```

Ogih, M., Okuyama, G., Otagawa, N., Oyeda, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojudoan, I., Rolfe, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I., Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Tang, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wellington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 161582)  
Morley, K.C.

Direct Submission  
Submitted (07-JAN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 10, 2001 this sequence version replaced gi:14600326.

Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc.help@bcm.tmc.edu](mailto:hgsc.help@bcm.tmc.edu)

Project Information  
Center project name: HMXT  
Center clone name: RP11-328H16

Summary Statistics  
Sequencing vector: M13; L08821  
Sequencing statistics: M77789  
Chemistry: Dye-terminator Big Dye; 2% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 176266 bases at least Q40  
Consensus quality: 181503 bases at least Q30  
Consensus quality: 185315 bases at least Q20  
Estimated insert size: 165934; sum-of-coverage estimation  
Quality coverage: 0x in Q20 bases; agarose-gel estimation  
Quality coverage: 9x in Q20 bases; sum-of-coverage estimation

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drift\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html)).  
NOTE: This is a working draft sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1  
83883: contig of 83883 bp in length  
83984: gap of unknown length  
83984: contig of 39777 bp in length  
123761: gap of unknown length  
123861: contig of 19596 bp in length  
143457: gap of unknown length  
143556: gap of unknown length  
149479: contig of 5922 bp in length  
149578: gap of unknown length  
149579: contig of 7172 bp in length  
156751: gap of unknown length  
156850: contig of 2551 bp in length  
159401: gap of unknown length  
159501: gap of unknown length  
159502: contig of 2081 bp in length.

Location/Qualifiers  
1. 161582  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-328H16"

BASE COUNT  
52133 a 29064 c 28547 g 51197 t 641 others

ORIGIN

Alignment Scores:  
Pred. No.: 2.27e+03  
Score: 161582  
Percent Similarity: 44.00  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 2  
Gaps: 0

US-09-698-781-17 (1-9) x AC020628 (1-161582)  
Qy 1 ThirleupheprovalleuLeupheleu 9  
Db 156549 ACTTTGTCACAGTCGCTTTCTTA 156575

RESULT 13  
AC090046/c  
LOCUS  
DEFINITION  
MUS musculus chromosome 6 clone RP23-82P24 map 6, WORKING DRAFT  
SEQUENCE, 12 unordered pieces.  
ACCESSION  
AC090046  
VERSION  
AC090046.3 GI:14336487  
KEYWORDS  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE  
MUS musculus.  
ORGANISM  
MUS musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
1 (bases 1 to 177387)  
Bairren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Batra, N., Bascien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Canarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P., Dearellano, R., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardaya, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karats, A., Larocque, K., Lamazares, R., Landers, T., Lenocsky, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mihova, T., Mlepta, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Ribbeck, M., Riley, R., Risse, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Soungue, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Teafaye, S., Theodore, J., Travers, M., Travis, J., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye-W, J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (11-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 11, 2001 this sequence version replaced gi:13357348.  
All repeats were identified using RepeatMasker:  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
Smit, A.F.A. & Green, P. (1996-1997)  
Center: Whitehead Institute/MIT Center for Genome Research  
Center code: MIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
Project Information  
Center project name: 82\_P24  
Center clone name: 112519  
Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 173803 bases at least Q40

```

Consensus quality: 175369 bases at least Q30
Consensus quality: 175928 bases at least Q20
Insert size: 180000; agarose-fp
Insert size: 176287; sum-of-contigs
Quality coverage: 10.5 in Q20 bases; agarose-fp
Quality coverage: 10.7 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 31880: contig of 31880 bp in length
*
31881 31980: gap of 100 bp
*
31981 32430: contig of 450 bp in length
*
32431 32530: gap of 100 bp
*
32531 33945: contig of 1415 bp in length
*
33946 34045: gap of 100 bp
*
34046 35388: contig of 1343 bp in length
*
35389 35488: gap of 100 bp
*
35489 41357: contig of 5869 bp in length
*
41358 41457: gap of 100 bp
*
41458 49412: contig of 7955 bp in length
*
49413 49512: gap of 100 bp
*
49513 58396: contig of 8884 bp in length
*
58397 58496: gap of 100 bp
*
58497 67955: contig of 9459 bp in length
*
67956 68055: gap of 100 bp
*
68056 115552: contig of 47497 bp in length
*
115553 115652: gap of 100 bp
*
115653 139731: contig of 24079 bp in length
*
139732 139831: gap of 100 bp
*
139832 162599: contig of 22768 bp in length
*
162600 162699: gap of 100 bp
*
162700 177387: contig of 14688 bp in length.
-----
FEATURES
Source
1. 177387
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="6"
/clone="RP23-82P24"
/clone_lib="RPCI-23 Female Mouse BAC"
1. 31880
/note="assembly-fragment"
clone_end:SP6
vector_side:left"
31981..32430
/note="assembly-fragment"
32531..33945
/note="assembly-fragment"
34046..35388
/note="assembly-fragment"
35489..41357
/note="assembly-fragment"
41458..49412
/note="assembly-fragment"
49513..58396
/note="assembly-fragment"
58497..67955
/note="assembly-fragment"
68056..115552
/note="assembly-fragment"
115653..139731
/note="assembly-fragment"
139832..162599
/note="assembly-fragment"
162700..177387
/note="assembly-fragment"
clone_end:T7

BASE COUNT 48909 a 39454 c 38603 g 49318 t 1103 others
ORIGIN
Alignment Scores:
Pred. NO.: 2.49e+03 Length: 177387
Score: 44.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-09-698-781-17 (1-9) x AC090046 (1-177387)
Qy 1 ThreupheProValleuPhelu 9
Db 127402 ACCCTTTCTGCTCTTCTTCTTA 127376
RESULT 14
AC121977
LOCUS
DEFINITION
Mus musculus chromosome UNK clone RP24-287A16, WORKING DRAFT
ACCESSION
AC121977
VERSION
AC121977.1 GI:21040100
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Mus musculus
ORGANISM
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 229084)
McPherson,J.D. and Waterston,R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 229084)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BB0287A16
----- Summary Statistics -----
Sequencing vector: M13: 0%
Sequencing vector: plasmid: 100%
Chemistry: Dye-terminator ET: 0% of reads
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 226270 bases at least Q40
Consensus quality: 226603 bases at least Q30
Consensus quality: 226772 bases at least Q20
Insert size: 222000; agarose-fp
Insert size: 228785; sum-of-contigs
Quality coverage: 14.73 in Q20 bases; agarose-fp
Quality coverage: 10.44 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 5993: contig of 5993 bp in length
*
5994 6093: gap of unknown length
*
6094 13020: contig of 6927 bp in length

```

BASE COUNT	ORIGIN
64105	a 47883 c 49570 g 66736 t 730 others
misc_feature	1..5993
/note="assembly_name:Contig7"	
misc_feature	6034..13020
/note="assembly_name:Contig8"	
misc_feature	13121..30634
/note="assembly_name:Contig9"	
misc_feature	30735..54503
/note="assembly_name:Contig10"	
misc_feature	54606..77711
/note="assembly_name:Contig11"	
misc_feature	77812..107448
/note="assembly_name:Contig12"	
misc_feature	107549..158557
/note="assembly_name:Contig13"	
misc_feature	158658..3229084
/note="assembly_name:Contig14"	

Alignment Scores:	
Pred. No.:	3.21e+03
Score:	44.00
Percent Similarity:	100.00%
Best local Similarity:	100.00%
Query Match:	100.00%
DB:	2
	Gaps: 0
US-09-698-781-17 (1-9) x AC121977 (1-229084)	Length: 229084
	Matches: 9
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

Qy	1	ThrLeupheProValLeuLeupheLeu	9
Db	89979	ACCCCTTTTCCTGTCCTTCTATTCCTA	90005

AC024305/c	AC024305	58558 bp	DNA	linear	HTG 13-JUL-2000
LOCUS					
DEFINITION	Homo sapiens clone RP11-24L17, LOW-PASS SEQUENCE SAMPLING.				

VERSION AC024305.1 GI:7108039  
KEYWORDS HTG; HTGS\_PHASE0.  
SOURCE Homo sapiens.

REFERENCE  
AUTHORS

Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 58558)  
Birren, B., Linton, L., Nussbaum, C. and Lander, E.

JOURNAL  
REFERENCE  
AUTHORS

Unpublished  
2 (bases 1 to 58558)  
Bitren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.

Boukhagelte, B., Brown, A., Burrett, G., Campopiano, A., Castle, A.,  
Chopelli, Y., Colangelo, M., Collins, S., Collimore, A., Cooke, P.,  
Dearellano, J., Dewar, K., Diaz, J. S., Dodge, S., Domino, M., Doyle, M.,  
Fenster, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D.,

TITLE  
JOURNAL  
COMMENT

Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J. C., Illey, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Lagocque, K., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McIvor, A., Mcneman, K., Mcpheeters, R., Meldrum, J., Menus, L., Milgrat, T., Miranda, C., Mlenga, V., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivar, T. M., Peterson, K., Pierre, N., Plesni, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, J., Thirrell, A., Subramanian, A., Talamas, J., Testayre, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: MIBR  
Web site: <http://www.seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information -----  
Center project name: L491  
Center clone name: 24\_L-17

```

NOTE: This record contains 60 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allow
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
892 991: gap of 100 bp in length
992 1884: contig of 893 bp in length
1885 1984: gap of 100 bp
1985 2670: contig of 866 bp in length
2671 2971: gap of 100 bp
2971 3853: contig of 883 bp in length
3854 3953: gap of 100 bp
4444: contig of 891 bp in length
4454 4944: gap of 100 bp
4945 5531: contig of 887 bp in length
5532 5931: gap of 100 bp
5932 6530: contig of 899 bp in length
6531 6930: gap of 100 bp
7019: contig of 889 bp in length
7220 7919: gap of 100 bp
7920 8797: contig of 878 bp in length
8798 8897: gap of 100 bp
8898 9778: contig of 881 bp in length
9779 9878: gap of 100 bp
9879 10789: contig of 911 bp in length
10790 10889: gap of 100 bp
11772: contig of 883 bp in length
11773 11872: gap of 100 bp
11873 12704: contig of 832 bp in length
12705 12804: gap of 100 bp
12805 13699: contig of 895 bp in length
13700 13799: gap of 100 bp
13800 14675: contig of 876 bp in length
14676 14775: gap of 100 bp
14776 15674: contig of 889 bp in length
15675 15774: gap of 100 bp

```

```

* 15775 16641: contig of 867 bp in length
* 16642 16741: gap of 100 bp
* 16742 17631: contig of 890 bp in length
* 17632 17731: gap of 100 bp
* 17732 18646: contig of 915 bp in length
* 18647 18746: gap of 100 bp
* 18747 19637: contig of 891 bp in length
* 19638 19737: gap of 100 bp
* 19738 20773: contig of 1036 bp in length
* 20774 20873: gap of 100 bp
* 20874 21733: contig of 860 bp in length
* 21734 21833: gap of 100 bp
* 21834 22695: contig of 862 bp in length
* 22696 22795: gap of 100 bp
* 22796 23622: contig of 827 bp in length
* 23623 23722: gap of 100 bp
* 23723 24605: contig of 883 bp in length
* 24606 24705: gap of 100 bp
* 24706 25574: contig of 869 bp in length
* 25575 25674: gap of 100 bp
* 25675 26520: contig of 846 bp in length
* 26521 26620: gap of 100 bp
* 26621 27494: contig of 874 bp in length
* 27495 27594: gap of 100 bp
* 27595 28461: contig of 867 bp in length
* 28462 28561: gap of 100 bp
* 28562 29445: contig of 884 bp in length
* 29446 29545: gap of 100 bp
* 29546 30421: contig of 876 bp in length
* 30422 30521: gap of 100 bp
* 30522 31380: contig of 859 bp in length
* 31381 31480: gap of 100 bp
* 31481 32375: contig of 895 bp in length
* 32376 32475: gap of 100 bp
* 32476 33345: contig of 870 bp in length
* 33346 33445: gap of 100 bp
* 33446 34280: contig of 835 bp in length
* 34281 34380: gap of 100 bp
* 34381 35230: contig of 850 bp in length
* 35231 35330: gap of 100 bp
* 35331 36198: contig of 868 bp in length
* 36199 36298: gap of 100 bp
* 36299 37169: contig of 871 bp in length
* 37170 37269: gap of 100 bp
* 37270 38133: contig of 864 bp in length
* 38134 38233: gap of 100 bp
* 38234 39116: contig of 883 bp in length
* 39117 39216: gap of 100 bp
* 39217 40103: contig of 887 bp in length
* 40104 40203: gap of 100 bp
* 40204 41074: contig of 871 bp in length
* 41075 41174: gap of 100 bp
* 41175 42077: contig of 903 bp in length
* 42078 42177: gap of 100 bp
* 42178 43010: contig of 833 bp in length
* 43011 43110: gap of 100 bp
* 43111 43994: contig of 884 bp in length
* 43995 44094: gap of 100 bp
* 44095 44965: contig of 871 bp in length
* 44966 45065: gap of 100 bp
* 45066 45939: contig of 874 bp in length
* 45940 46039: gap of 100 bp
* 46040 46928: contig of 889 bp in length
* 46929 47028: gap of 100 bp
* 47029 47881: contig of 853 bp in length
* 47882 47981: gap of 100 bp
* 47982 48855: contig of 874 bp in length
* 48856 48955: gap of 100 bp
* 48956 49850: contig of 895 bp in length
* 49851 49950: gap of 100 bp
* 49951 50825: contig of 875 bp in length
* 50826 50925: gap of 100 bp
* 50926 51781: contig of 856 bp in length

```

```

* 51782 51881: gap of 100 bp
* 51882 52673: contig of 792 bp in length
* 52674 52773: gap of 100 bp
* 52774 53716: contig of 943 bp in length
* 53717 53816: gap of 100 bp
* 53817 54684: contig of 868 bp in length
* 54685 54784: gap of 100 bp
* 54785 55646: contig of 862 bp in length
* 55647 55746: gap of 100 bp
* 55747 56615: contig of 869 bp in length
* 56616 56715: gap of 100 bp
* 56716 57586: contig of 871 bp in length
* 57587 57686: gap of 100 bp
* 57687 58558: contig of 872 bp in length.
*
FEATURES
  source          1..58558
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="RP11-24L17"
                  /clone_lib="RPC1-11 Human Male BAC"
BASE COUNT      15372 a 10594 c 10818 g 15552 t      6222 others
ORIGIN
Alignment Scores:
Pred. No.:      1.3e+03      Length:      58558
Score:          43.00       Matches:      8
Percent Similarity: 100.00%   Conservative: 1
Best Local Similarity: 88.89%   Mismatches:  0
Query Match:    97.73%       Indels:      0
DB:             2           Gaps:        0
US-09-698-781-17 (1-9) x AC024305 (1-58558)
QY      1 ThrLeuPheProValLeuLeuPheLeu 9
Db      30338 ACCCTTTTCCATATTATTATTCTT 30312

```

Search completed: March 14, 2003, 04:41:17  
Job time : 213.888 secs